John Beavers, The University of Texas at Austin
jtbeavers@utexas.edu

Seoul National University
Sep 18th, 2020

Event Structures and The Semantic Typology of Verbal Roots Across Languages

(Joint work with Michael Everdell*, Kyle Jerro†, Henri Kauhanen‡, Andrew Koontz- Garboden§, Elise LeBovidge|, and Stephen Nichols§
∗ = The University of Texas at Austin, † = University of Essex, ‡ = Zukunftskolleg, University of Konstanz, § = University of Manchester, | = University of Washington)

1 Introduction

- Among the most fundamental questions about verb meanings are what their basic building blocks are and how those building blocks are composed into more complex meanings. One key insight is that the answer involves a decomposition into basic event types (Lakoff 1965):

  (1) a. John flattened the rug ≈ John **caused** the rug to **become** flat.
  b. John cracked the vase ≈ John **caused** the vase to **become** cracked.

- It is thus now often assumed that verb meanings consist partly of an "event structure" built from (a) a template of basic eventive predicates (e.g. *v*s; Marantz 1997) and (b) idiosyncratic roots filling in real world meanings (e.g. manner, state; Rappaport Hovav and Levin 1998):

  (2) a. John flattened the rug ≈ [$_{vP}$ John [$_{v'}$ v_{cause} [$_{vP}$ the rug [*-en*$_{v_{\text{become}}}$ $\sqrt{\text{FLAT}}$]]]]
  b. Kim cracked a vase ≈ [$_{vP}$ Kim [$_{v'}$ v_{cause} [$_{vP}$ a vase [$_{v'}$ v_{become} $\sqrt{\text{CRACK}}$]]]]

- The template defines the verb's lexical aspectual properties, argument structure, and regular derivational morphology; the root just determines the verb's idiosyncratic morphology.

- An underexplored question (though see Dowty 1979, Goldberg 1995, Wechsler 2005, a.o., for related discussion) is whether there is an equivalent clean divide between meanings entailed by roots and by templates, e.g. are cause and become *only* introduced templatically?

- This is Embick's (2009) "Bifurcation Thesis for Roots" (BTR), Arad's (2005) "Root Hypothesis" (also Borer 2005, Dunbar and Wellwood 2016), and it makes a strong prediction: entailing templatic meaning requires templatic syntax!

- But is this true? Answering this is important for shoring up event structural theories. If roots can mean anything then it potentially undermines what kind of predictions we can make about the syntax/semantics correlation (as per the dire warning of Dowty 1979: 125-126).

- Beavers and Koontz-Garboden (2020) propose that some English roots entail templatic meaning, and have grammatical effects. However, these results could be flukes of English.

- We present a broad cross-linguistic study into this question, focusing on the change entailment in change of state verbs, a meaning that on all approaches is templatic. We claim that:

  – Certain classes of roots of change-of-state verbs entail change regardless of the template they occur in across languages, suggesting that the root lexically entails change.
  – We show with typological evidence that the same root classes have morphological properties distinct from other change-of-state verb roots.

- Nonetheless, we show that this does not mean there are no predictions to be had from event structural theories, just that they are more complicated than previously expected.

2 The Roots of Canonical Change-of-State Verbs

- The most common understanding of (2) according to the BTR is that the root is purely stative; the entailment of change comes from v_{become}. All roots that occur in this template should be semantically and grammatically identical in their behavior, modulo lexical idiosyncrasy or subregularity orthogonal to the BTR (e.g. phonological or historical effects).

- But there are significant divergences from this (cf. Megerdoomian 2002, Koontz-Garboden 2006). Beavers and Koontz-Garboden (2020) distinguish Dixon's (1982) **property concept (PC) roots** (e.g. Levin's 1993: 245 deadjectival change-of-state verbs) from **result roots** (e.g. Levin's non-deadjectival change-of-state verbs). The roots for our study are:

  (3) **Property Concept** (deadjectival verbs)

  a. *Dimension*: large/big/enlarge, small/shrink/shrunken, short/shorten, long/lengthen, deep/deepen, wide/widen, tall/height/heighten

  b. *Age*: old/aged/age

  c. *Value*: bad/worsen/worse, good/improve/improved

  d. *Color*: white/whiten, black/blacken, red/redden, green/make green, blue/make blue, brown/make brown

  e. *Physical Property*: cool/cool, cold/make cold, warm/warm, hot/heat up, dirty/dirty, dry/dry, wet/wetted, straight/straighten, hard/harden (tough/toughen), soft/soften, tight/tighten, clear/clear, clean/clean, smooth/smooth, sharp/sharpen, sweet/sweeten, weak/weaken, strong/strengthen

  f. *Speed*: fast/speed up, slow/slow down

  (4) **Result Roots** (non-deadjectival verbs)

  a. *Entity-specific Change of State*: burned/burn, melted/melt, frozen/freeze, decayed/decay (rotten/rot), swollen/swell, grown/grow, bloomed/bloom (flowered/flower, blossomed/blossom), withered/wither (wilted/wilt), fermented/ferment, sprouted/sprout (germinated/germinate), rusted/rust, tarnished/tarnish

  b. *Cooking Verbs*: cooked/cook (baked/bake, fried/fry, roasted/roast, steamed/steam), boiled/boil

  c. *Breaking Verbs*: broken/break, cracked/crack, crushed/crush, shattered/shatter, split/split, torn/tear (ripped/rip), snapped/snap

  d. *Bending Verbs*: bent/bend, folded/fold, wrinkled/wrinkle, creased/crease

  e. *Verbs of Killing*: dead/killed/kill, murdered/murder, drowned/drown

  f. *Destroying Verbs*: destroyed/destroy (ruined/ruin)

  g. *Verbs of Calibratable Change of State*: differ/different, go up (raised/rise, ascended/ascend, increased/increase, gained/gain), go down (fallen/fall, dropped/drop, descended/descend, decreased/decrease, declined/decline)

  h. *Verbs of Inherently Directed Motion*: come/came, gone/go, go in (entered/enter), go out (exited/exit), returned/return

- We first recap and expand Beavers and Koontz-Garboden's argument that English result roots show morphological and semantic properties that violate the BTR. We suggest a somewhat principled analysis of this, albeit that amounts to a series of language particular stipulations.

- Beavers and Koontz-Garboden suggest an alternative that relies on a more fundamental assumption violating the BTR at its core, namely that result roots entail change. Here we present cross-linguistic evidence argues for this analysis over the language-particular one.

3 Morphological and Semantic observations

#1 One prediction of the BTR is that barring lexical idiosyncrasy or BTR-orthogonal (e.g. phonological) subregularity, all change-of-state verb roots should show all of the same forms.

- English PC roots have two stative forms: a simple adjective and a deverbal *–en* one.

  (5) a. Look at the bright picture on your left. (=camera took a bright picture)
  b. Look at the brightened picture on your left. (e.g. a digitally brightened picture)

- Embick (2004) analyzes these as the same root occurring in two adjectivalizing contexts: just a root with adjectivalizing Asp head, or a verbal structure also including v_{become}:

  (6) a. Basic states (cp. Embick 2004: 363): [$_{\text{AspP}}$ Asp $\sqrt{\text{ROOT}}$]
  b. Result states (cp. Embick 2004: 367): [$_{\text{AspP}}$ Asp$_R$ [$_{v\text{P}}$ DP v_{become} $\sqrt{\text{ROOT}}$]]

- Crucially, with result roots there is just one morphological form, the *–en* form.

  (7) broken, chipped, cracked, crashed, crushed, fractured, ripped, shattered, smashed, snapped, splintered, split, torn, baked, barbecued, blanched, boiled, braised, ...

- Under the BTR, any stative root should appear in either of (6), yet result roots seem to not appear in (6a). Embick (2004: 358) claims they do, but with them *-ed/en* realizes both Asp and Asp$_R$, while with PC roots Asp$_R$ is *-ed/en* and Asp null, an accident of English.

#2 However, this makes predictions about meanings of the various stative forms:

  (8) a. Simple adjectives from PC roots (e.g. *red*) will not entail prior change.
  b. Derived adjectives from PC roots (e.g. *reddened*) will entail prior change.
  c. Result root adjectives will not entail prior change (owing to the (6a) reading).

- (8a,b) are borne out, but (8c) is not (Koontz-Garboden 2005, 2010, Deo et al. 2011).

  (9) a. The bright/#brightened photo has never (been) brightened.
  b. The long/#lengthened river has never (been) lengthened.
  c. The red/#reddened dirt has never (been) reddened.
  (10) a. #The shattered vase has never (been) shattered.
  b. #The dead man never died.
  c. #The cooked chicken has never (been) cooked.

- Perhaps we haven't clarified what the relevant simple state would be. However, even taking the simple state to be the "prototypical" one (the outcome of an event had it occurred), we don't get the right reading, e.g. (10a) doesn't improve if one makes pieces that form a vase.

- NB: there are the "derived statives" of Nedjalkov and Jaxontov (1988), but there is still a change, albeit atemporal: it is measured along a spatial scale (Gawron 2009, Koontz-Garboden 2010) (see Deo et al. 2011, 2013 on non-spatial atemporal change).

  (11) ...this paper provides a guide for writing letters that extend below the baseline. Internal **broken** lines serve as a reference for writing half-space letters.
  (`http://www.abledata.com/abledata.cfm?pageid=19327&top=11104&ksectionid=0`

- Also, there is cross-speaker variation: for some speakers *broken* has a "non-functioning" reading not requiring prior change (cp. *closed*, which does not require prior change either). But all that matters for us is that *some* result root adjectives require change.

#3 Maybe English result roots require a v_{become} (cf. Pross 2016). But we also see a split under *again*. PC root verbs allow restitutive and repetitive readings, a presumably scopal effect:

(12) John flattened the rug again.

a. [$_{v\text{P}}$ John [$_{v'}$ v_{cause} [$_{v\text{P}}$ the rug [-*en*$_{v_{become}}$ [$\sqrt{\text{FLAT}}$ again]]]]] (restitutive)

b. [[$_{v\text{P}}$ John [$_{v'}$ v_{cause} [$_{v\text{P}}$ the rug [-*en*$_{v_{become}}$ $\sqrt{\text{FLAT}}$]]]] again] (repetitive)

- Positing an obligatory v_{become} will leave the root exposed to restitutive modification. Yet result roots lack restitutive readings (Rappaport Hovav 2010: 7, Beavers and Koontz-Garboden 2012: 358), even in contexts that should favor a one (as best as one can tell):

(13) [John kills a rabbit, takes it home and skins and butchers it and then puts the fresh meat in the freezer for a week. He then takes it out and puts it on the table to thaw.]
#John thawed the meat again. (necessarily two defrostings)

(14) [A store makes their shirts in the back. John buys one and leaves with it, but then decides he does not want it. He takes the shirt back to exchange it.]
#John returned the shirt again. (necessarily two returnings)

(15) [John lives in a hot region and finds a fruit with brown, fatty edges. He takes it home, trims off the edges, and puts it in the fridge. He later takes it out and fries it.]
#John fried the fruit again. (necessarily two fryings)

- Maybe result roots resist sublexical modification, forcing high attachment, but why?

#4 English PC root verbs tend to be marked (barring phonology) while result root are unmarked:

(16) a. wid**en**, whit**en**, straight**en**, stiff**en**, short**en**, **en**large, hard**en**, etc.

b. burn, melt, freeze, cook, break, crack, crush, shatter, murder, wrinkle, etc.

- We might try the same sort of analysis we tried for the result roots, and assume that PC roots come with a feature forcing them to combine with Asp first, permitting (17b) and not (17a).

(17) a. Basic verbs: [$_{v\text{P}}$ v_{become} $\sqrt{\text{ROOT}}$] (result roots)

b. Derived verbs: [$_{v\text{P}}$ v_{become} [$_{\text{AspP}}$ Asp $\sqrt{\text{ROOT}}$]] (PC roots)

- Descriptively, PC roots lexicalize as adjectives and result roots as verbs. We could even posit that *all* roots resist *again* modification, forcing attachment at Asp for PC roots and v_{become} for result roots, deriving the right semantics.

(18) a. [$_{v\text{P}}$ John [$_{v'}$ v_{cause} [$_{v\text{P}}$ the rug [-*en*$_{v_{become}}$ [$_{\text{AspP}}$ [$_{\text{AspP}}$ Asp $\sqrt{\text{FLAT}}$] again]]]]]

b. [$_{v\text{P}}$ John [$_{v'}$ v_{cause} [$_{v\text{P}}$ the rug [$_{v'}$ [$_{v'}$ -*en*$_{v_{become}}$ $\sqrt{\text{FLAT}}$] again]]]]

- But we now we have a coincidence of three separate *syntactic* assumptions — two arbitrary divisions of category features on roots plus one arbitrary constraint on sublexical modification. Is there a deeper explanation?

- An emergent generalization is that the states described by result roots are never dissociable from change — when you use a root, you get change. Maybe change is *part of the state itself*, and thus by naming such a state the root entails change as well, violating the BTR.

(19) a. $[\![\sqrt{\text{FLAT}}]\!] = \lambda x \lambda s[flat'(x, s)]$
b. $[\![\sqrt{\text{CRACK}}]\!] = \lambda x \lambda s[cracked'(x, s)]$, where $\forall x \forall x[cracked'(x, s) \rightarrow \exists e'[become(e', s)]]$

- We no longer need the syntactic features to get the semantics. Assuming cross-cultural similarity, we'd also expect the semantic patterns to recur crosslinguistically, a prediction we do *not* get if it's a syntactic fact of English. We explore this next.

4 Cross-Linguistic Semantic Study of Change-of-State Verbs

- To test the semantic predictions we did native speaker studies of a small set of translations of English terms as a proof-of-concept, using Greek (Spathas 2017), Kinyarwanda (Northeastern Bantu) (Jerro 2017), Kakataibo (Panoan) (Valle et al. 2017), Marathi (Indic, Indo-European), and Hebrew (Semitic, Afro-Asiatic). I present Kakataibo as a case study.

- Kakataibo simple states and inchoatives are usually labile, though the former are adjectives and the latter verbs, while causatives are generally formed by causative *-o* (plus some anti-causative and equipollent relationships), and result states by factive *-kë*:[1]

(20)

Language	Root	simple state	inchoative	causative	result state
Kakataibo	large	*ani*	*ani*	*ani-o*	*ani-kë/ani-o-kë*
Kakataibo	wrinkle	—	*churi*	*churi-o*	*churi-kë/churi-o-kë*

- We examined the following culturally appropriate terms:

(21) **PC**: *xo paxada* 'yellow', *ani* 'big', *upi(t)* 'pretty, beautiful', *tuna(n)* 'black', *uxu(a)* 'white', *inru* 'hard', *xana* 'hot', *bata* 'sweet', *bachu* 'soft', *ˈaidama* 'bad', *bënsi(t)* 'thin', *xëni* 'old', *chadkë(t)* 'thin', *chabat* 'wet', *puntë(t)* 'straight', *diba(t)* 'smooth', *kacha* 'sour', *anaha* 'wide', *chukúma* 'small', *cha* 'big', *mamúa* 'round', *ëd-ki-kë* 'dry-INTR-NMLZ'

(22) **Result**: *tëa-kë* 'cut-NMLZ', *nën* 'burn', *a-ru-kë* 'do-UP-NMLZ', *tun-ka-kë* 'shoot-TR-NMLZ', *rëtë* 'kill/murder', *këñu* 'exterminate', *chachi* 'stab', *xui* 'barbecue', *sasa-ka* 'fry-TR', *musa* 'mix/stir/mash'

- In general PC vs. result roots seem to show a distinction regarding change entailments. Simple PC statives do not entail change but derived PC statives do:

(23) a. *báinka* ***ani*** *ˈikë* *ˈaibika* *uini* *abi* *ni*
báin=ka=a ani ˈikë ˈai=bi=ka=a uini a=bi ni
hill=VAL=3A/S big be.3.IMPF then=EMPH=VAL=3A/S INDF.PRO 3=EMPH nor
Diosabi *ni* *uni* *yubë* *unibi* *anioima.*
Diosabi ni uni yubët uni=bi ani-o-i-i=ma.
God=EMPH nor man sorcerer man=EMPH big-FACT-IMPF=PROX=NEG
'The hill is big, but nobody nor God nor a sorcerer made it big.'

[1] 1-10(S)=noun classes, 3=third person, A=subject of transitive verb, EMPH=emphatic, FACT=factitive, FV=final vowel, IMPF=imperfective, INDF=indefinite, INTR=intransitive, ITR=iterative, NEG=negation, NFUT.NMLZ=non=future nominalizer, PASS=passive, PAST=past, PFV=perfective, POSS=possessive, PRO=pronoun, PRES=present, PROX=proximate, S=subject of intransitive verb, UP=up, VAL=validational.

b. *#taíka* ***puntëokë*** *ikë* *aibika*
tain=ka=a puntët-o-kë ikë ai=bi=ka=a
arrow.stick=VAL=3A/S straight-FACT-NFUT.NMLZ be.IMPF.3 then=EMPH=VAL=3A/S
uini *abi* *puntëoima* *iáxa.*
uini a=bi puntët-o-i=ma i-a-x-a
INDF.PRO 3=EMPH straight-FACT-A/S>S:SE=NEG be-PRFV-3-N.PROX
'The tree (used to make arrows) stem is straightened but nobody made it straight.

- Statives of result roots pattern like derived statives of PC roots exclusively:

(24) *#nū* *nami* ***tëakë*** *ˈikë* *ˈaibika* *uini* *abi*
nū nami tëa-kë ˈikë ˈaibika uini abi
thing flesh cut=NFUT.NMLZ be.3.IMPF then=EMPH=VAL=3A/S INDF.PRO 3=EMPH
tëakëma *ˈikë.*
tëa-kë=ma ˈikë.
cut=NFUT.NMLZ=NEG be.3.IMPF
'The meat is cut but nobody cut it.'

- PC roots allow restitutive readings under iterative *-tëkën* marking, while result roots usually resist them (though it was hard to come up with what the simple state would be):

(25) a. [The desert starts off dry. Then, it is made non-dry. Then it turns dry again.]
madin *papanka* ***ëdkitëkënia***.
madi=n papa=n=ka=a ëd-ki-tëkën-i-a.
sand=POSS father=A/S=VAL=3A/S dry-INTR-AGAIN-IMPF=N.PROX
'The desert is getting dry again.'

b. [The man picks up a banana. A wizard makes it inedible. The man fries it.]
#uninka *nodi* ***sasakatëkënia***.
uni=n=ka=a nodi sasa-ka-tëkën-a-x-a/'a-ru-tëkën-a-x-a
man=A/S=VAL=3A/S banana fry-TR-AGAIN-PFV-3-N.PROX
'The man fried the banana again.'

- There was some variation from English, e.g. Kakataibo *kill* allowed restitutive modification.

(26) [The stone was always dead. Then, it was brought to life. Then, I kill it.]
maxákana ***rë(të)tëkëa***.
maxat=ka=na rëtë-tëkën-a
stone=VAL=1A/S kill-AGAIN-PFV
'I killed the stone again.' (≈ I made the stone be not alive again)

- This is expected, however: we do no expect all translation equivalents to be perfectly equivalent. In turn, sometimes PC roots seemed to generate an inference of change:

(27) *#fierro* ***inru*** *ikë* *aibika* *inrumi* *ama* *ikë.*
fierro inru ikë aibika inrumi ama ikë
fierro hard be.IMPF.3 then=EMPH=VAL=3A/S hard.CAUS.A/S>S:SE 3=NEG be.IMPF.3
'Construction metal is hard but nobody made it hard.

- However, *fierro* is a man-made metal, and this may be the source of the contradiction.

Upshot: While PC roots behave as predicted by the BTR having a purely stative meaning, result roots can have an entailment of change, a fact that holds up in several languages.

5 Typological Study on the Morphology of Change-of-State Verb Roots

5.1 Basic Methodology

- To test the morphological predictions we examined a balanced language sample for various PC and result root meaning forms, looking at their stative and eventive uses.
- We targeted the WALS 100 Language list (Dryer and Haspelmath 2013). If we lacked sufficient grammatical and dictionary resources we substituted languages from the WALS 200 list, and also added in a few based on available resources or native speakers/fieldworkers for a total of 88 languages, mostly covering the areas and families of WALS 100 (see §A).
- We used the 72 English adjectives/verbs in (3) and (4) (or appropriate source language equivalents). For 11 languages we had native speakers who helped us with the paradigms. For the rest, we did bidirectional dictionary searches. We did not assume all translations were perfect, just similar enough to justify being in the same broad/medium-grained semantic class.
- We collected paradigms of simple state-inchoative-causative-result state (e.g. *red-redden-redden-reddened*), plus also any underlying root for languages utilizing bound roots:

(28)

Language	**Root**	**underlying root**	**simple state**	**inchoative**	**causative**	**result state**
Tzeltal	small	—	*tut*	*tut-ub*	*tut-ub-tes*	*tut-ub-en*
Oromo	long	*dheer-*	*dheer-aa*	*dheer-addh*	*dheer-essuu*	—

- We privileged lexicalized terms over derived terms, and morphologically over periphrastically derived terms, motivated by a general tendency across languages for more lexical(ized) forms to be "normal" or "default" expressions (see e.g. Shibatani 1976, Shibatani and Pardeshi 2001, Harley 2008). An example complete paradigm is given in §B for Kinyarwanda.
- If a form was unattested but our resources gave productive processes for deriving it. This was strong evidence that at least *some* forms existed, though we don't know which. But it could affect our results. This was especially problematic for highly agglutinating languages such as Kiowa where some dictionaries give stems and rules rather than full lists of forms (Watkins and McKenzie 1984: 153.
- To check for this we constructed hypothetical forms (marked by @):

(29)

Language	**Root**	**underlying root**	**simple state**	**inchoative**	**causative**	**result state**
Kiowa	big	—	*ét*	*@ét-ɔ́m-gyá*	*@ét-ɔ́m*	*@ét-dɔ́·*

- We do not include hypotheticals in the data discussed below, but all of our tests were run with hypotheticals included and nothing changed.
- But we also coded morphological relations as a 5-sequence code relating each form X to its underlying root, simple state, inchoative, causative, and result state in that order.

(30)

	position within form X's 5-sequence code				
	1st	**2nd**	**3rd**	**4th**	**5th**
show's X's relationship to the:	underlying root	simple state	inchoative	causative	result state

- Values for X related to Y are X is **l**abile with Y, **u**nrelated to Y, **e**quipollent with Y, **d**erived from Y, **i**nput to Y, **t**ransitively related to Y (via form Z), Y is X (**s**elf), or Y is **n**ot attested.

(31)

Language	Root	underlying root	simple state	inchoative	causative	result state
Tzeltal	small	—	*tut*	*tut-ub*	*tut-ub-tes*	*tut-ub-en*
		—	nsitt	ndsii	ntdse	ntdes
Oromo	long	*dheer-*	*dheer-aa*	*dheer-addh*	*dheer-essuu*	—
		siiin	dseen	desen	deesn	—

- We made the following assumptions in selecting the final data for the analyses below:
  - We ignored a root meaning in a language if we had zero data (i.e. it's a resource gap).
  - If we had several synonyms for a root meaning we selected one random synonym (a Monte Carlo simulation showed that over 1,000 trials synonym choice did not matter).
- We ended up with 3,368 PC and 3,500 result roots with data, totaling 6,868 paradigms and 34,340 possible forms. 15,169 were recorded (and there were 2,957 hypotheticals, ignored here). Once one synonym is chosen per synonym set there are 2,712 PC and 2,417 result roots with data, with the exact number of cells filled depending on which random synonyms were chosen. A searchable database is at https://verbal-roots.la.utexas.edu/.

5.2 The Existence of Simple Stative Forms

- PC roots overwhelmingly tend to have simple stative forms (that usually serve as input to the rest) and result roots overwhelmingly tend not to, as per §3. See §B for a typical example.
- We calculated for each root the percentage of languages for which we had a simple state form among those that had data for that root at all and compared the distributions (medians) for PC vs. result roots. The results in Fig. 1 were statistically significant (on a Mann Whitney U-test; PC = 95.67%, result = 1.59%, $U = 1266.5$, $n_1 = n_2 = 36$, $p < 0.001$ one-tailed).

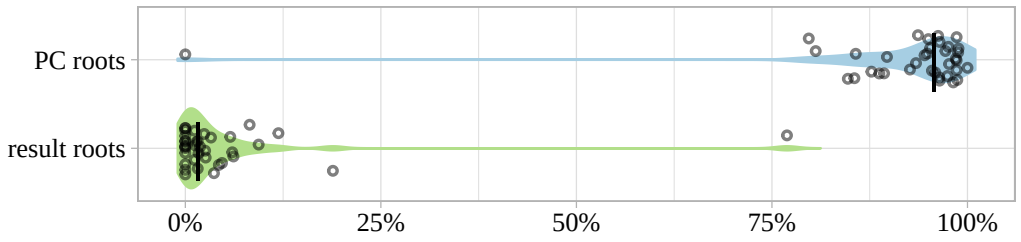

Figure 1: *Proportion of Languages With Underived States by Root Class Coded by Translation*

- Comparing the various subclasses produced some interesting results as in Fig. 2, with only two odd cases. Age is a PC root that patterns like a result root. This was a coding decision; one cannot be old without having started young, so we classed *old* and its translations result states. For calibratable change *differ* was a verb but *different* was coded as a simple state.

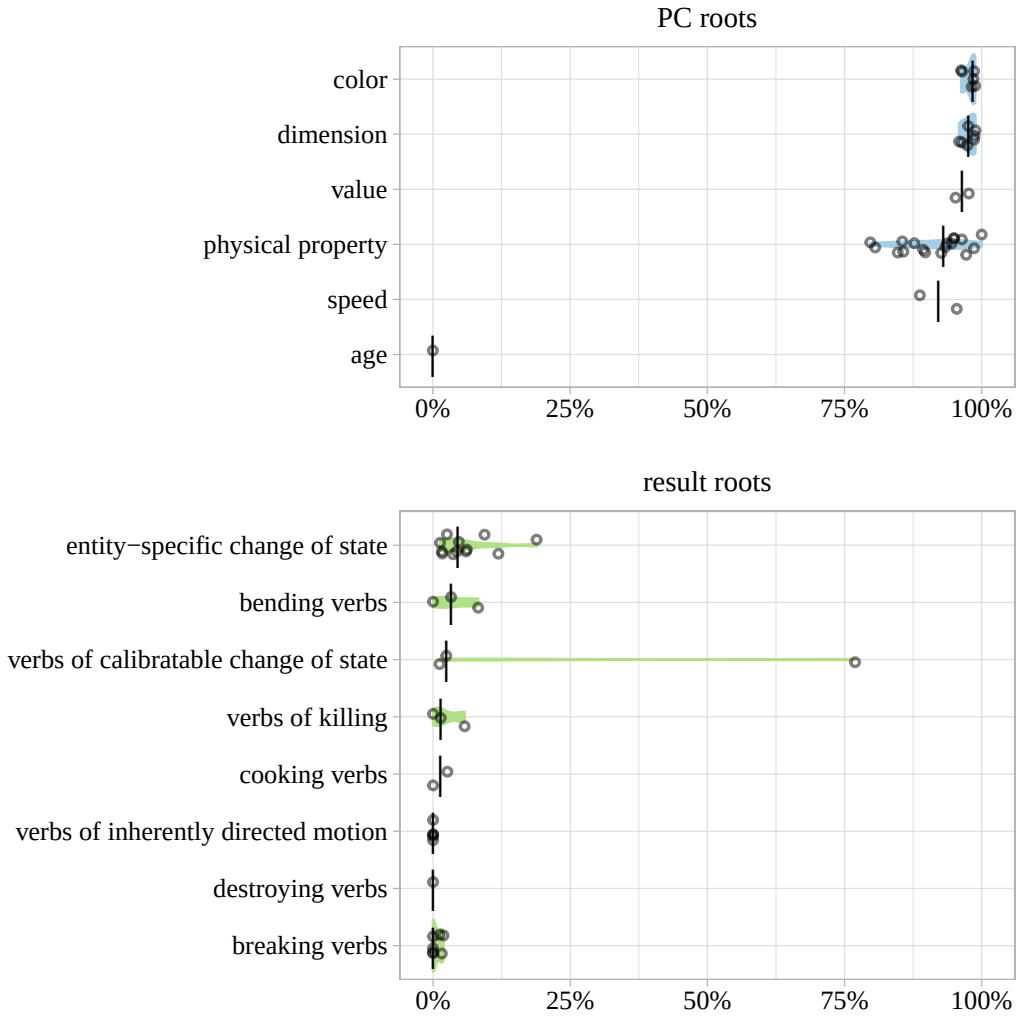

Figure 2: *Proportion of Languages With Underived States by Root Subclass Coded by Translation*

- A potential objection is that some terms might have semantically simple stative meanings but there is no corresponding English simple state form, so a deverbal result was used instead.
- However, we could code all statives as simple unless derived from a verb, using morphology as a proxy for translation bias. The results (Fig. 2) did not change in significance (PC median = 95.61%, result median = 27.66%, $U = 1294$, $n_1 = n_2 = 36$, $p < 0.001$ one-tailed).

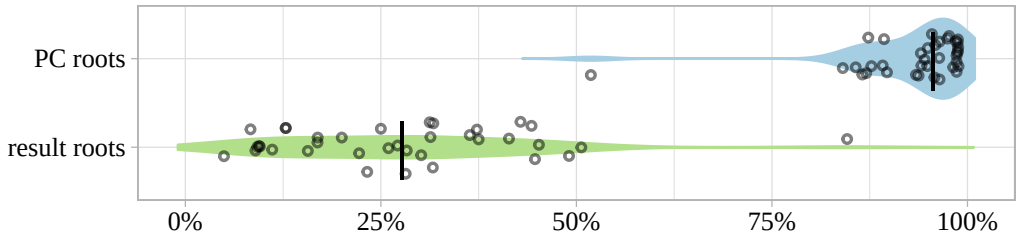

Figure 3: *Proportion of Languages With Underived States by Root Class Coded by Morphology*

- A second objection is that we have categorial translation bias: when looking for (de)verbal words we'll find (de)verbal words. But this is not so. First, Dixon (1982) showed categories for the "same" meanings can vary in category across languages. Second, our translation bias test showed that many deverbal translations are not obviously deverbal in other languages.

5.3 Preferences for Marked vs. Unmarked Verbal Forms

- We coded every causative and inchoative as "marked" iff it was overtly derived from or equipollent to something. A verbal paradigm is marked iff both verbal forms are marked.
- For each root we calculated the percent of languages with a marked verb paradigm and compared the distributions for PC vs. result roots. The results (Fig. 4) were significant (PC median = 56.01%, result median = 15.20%, $U = 1291$, $n_1 = n_2 = 36$, $p < 0.001$ one-tailed).

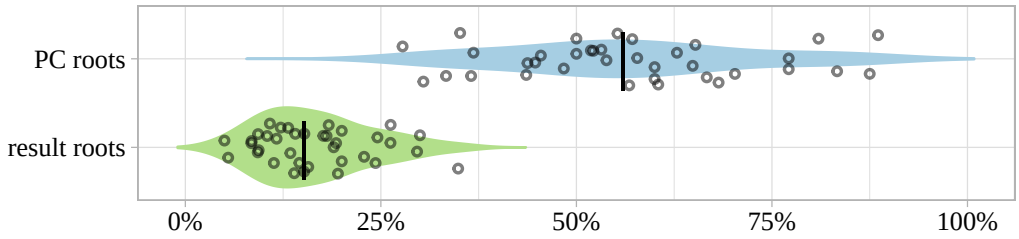

Figure 4: *Proportion of Languages With Marked Verbal Paradigms States by Root Class For Complete Paradigms*

- However, unlike the presence/absence of simple stative terms, here we don't have the same kind of floor/ceiling effect (e.g. PC root verbal paradigms aren't universally marked and result root paradigms aren't universally unmarked). Why?
- Some interesting typological trends emerge that explain this. Consider Fig. 5, where we rank languages by the relative amount of result root verb paradigm markedness over PC result root verb paradigm markedness.
- The expectation is that languages should show a ratio well below 1.0. This is indeed largely the case, but three additional and very telling cases emerge.

#1 Four languages (boldfaced in Fig. 5) show more verbal markedness among result roots than PC roots. There are reasons to not consider the four boldfaced languages exceptional.

- Paumarí is an equipollent language, where everything is derived from an underlying root:

(32)

language	**root**	**underlying root**	**simple state**	**inchoative**	**causative**	**result state**
Paumarí	tough	*dakha-*	*dakhaki*	*a'dakhaki*	*bina'dakhaki*	—

- We had only two PC paradigms, yet one involved suppletion (*sapasapaki* 'wide', *bavi bini'avini* 'widen'), skewing the results. Thus the problem is low data. A more natural view is that all paradigms are marked regardless of the PC vs. result distinction, save one irregular.
- Languages having less than 12 sufficiently complete verbal paradigms among PC or result roots (i.e. less than 33.33% of the minimum 36 each), are italicized in Fig. 5.

Language	# Verb paradigms PC	# Verb paradigms RR	% Markedness PC	% Markedness RR	% Markedness RR/PC	Language	# Verb paradigms PC	# Verb paradigms RR	% Markedness PC	% Markedness RR	% Markedness RR/PC
Kinyarwanda†	24	33	4.17	9.09	2.18	Greek (Modern)	22	44	59.09	6.82	0.12
Paumar*	2	10	50.00	100.00	2.00	Vietnamese	31	35	87.10	8.57	0.10
Navajo*	18	18	66.67	94.44	1.42	Tenango Tzeltal	34	40	76.47	7.50	0.10
Kakataibo†	59	64	23.73	31.25	1.32	Russian	37	40	83.78	7.50	0.09
Gjjolaay Eegimaa*	31	26	80.65	88.46	1.10	*Huitoto (Minica)**†	11	42	27.27	2.38	0.09
*Oneida**	9	11	44.44	45.45	1.02	*Martuthunira*	6	19	100.00	5.26	0.05
Burushaski†	16	49	6.25	6.12	0.98	Swahili	36	42	50.00	2.38	0.05
Hebrew (Modern)*	35	42	100.00	97.62	0.98	Yup'ik	14	30	71.43	3.33	0.05
Arabic (Egyptian)*	24	40	100.00	90.00	0.90	*Georgian*	9	34	66.67	2.94	0.04
Paiwan	14	20	92.86	80.00	0.86	English	43	60	46.51	1.67	0.04
Cree (Plains)*	16	52	62.50	50.00	0.80	Spanish	34	36	88.24	2.78	0.03
Acholi	24	40	70.83	55.00	0.78	*Alamblak*	2	10	50.00	0.00	0.00
Berber (Middle Atlas)*	26	44	73.08	54.55	0.75	Barasano†	19	31	10.53	0.00	0.00
Tagalog*	14	30	57.14	40.00	0.70	Burmese†	28	49	3.57	0.00	0.00
Yagua†	34	53	14.71	9.43	0.64	*Chamorro*†	11	34	9.09	0.00	0.00
Hausa	27	28	59.26	35.71	0.60	French	46	41	69.57	0.00	0.00
Malagasy	8	22	100.00	59.09	0.59	Khoekhoe†	12	53	8.33	0.00	0.00
Oromo (Harar)*	17	17	94.12	52.94	0.56	*Koasati*†	10	30	30.00	0.00	0.00
Finnish	56	48	96.43	54.17	0.56	*Meithei*	4	40	50.00	0.00	0.00
Pintupi	14	26	85.71	46.15	0.54	*Mixtec (Chalcatongo)*	10	38	60.00	0.00	0.00
Yaqui	11	38	45.45	23.68	0.52	Mocovi†	16	34	25.00	0.00	0.00
Persian	14	35	57.14	25.71	0.45	Otomi†	46	74	15.22	0.00	0.00
Lakhota†	10	24	10.00	4.17	0.42	Quechua (Huallaga)	18	44	38.89	0.00	0.00
Kwoma†	10	25	10.00	4.00	0.40	*Tiwi*†	9	20	11.11	0.00	0.00
Korean	16	45	43.75	15.56	0.36	Yoruba	40	97	40.00	0.00	0.00
Hindi	19	51	47.37	15.69	0.33	Zulu†	29	46	13.79	0.00	0.00
Japanese	18	45	83.33	26.67	0.32	Anejom̃†	17	35	0.00	0.00	—
Khalkha	34	26	85.29	26.92	0.32	Bariai†	12	39	0.00	0.00	—
Guaran†	13	43	7.69	2.33	0.30	*Chukchi*†	4	19	0.00	0.00	—
Kayardild	6	22	100.00	27.27	0.27	*Daga*†	5	23	0.00	4.35	—
Karok†	7	28	28.57	7.14	0.25	*Fijian*	0	19	—	5.26	—
Basque	27	35	70.37	17.14	0.24	*Gooniyandi*	0	10	—	0.00	—
Warao	14	23	71.43	17.39	0.24	Indonesian†	19	47	0.00	6.38	—
Carib	13	19	92.31	21.05	0.23	*Jakaltek*	0	24	—	0.00	—
Turkish	39	43	79.49	16.28	0.20	*Kannada*†	8	33	0.00	9.09	—
Mapudungun	27	25	40.74	8.00	0.20	*Kewa*†	5	34	0.00	0.00	—
Tsimshian (Coast)†	7	37	14.29	2.70	0.19	*Kiowa*†	1	8	0.00	0.00	—
Dani (Lower Grand Valley)	2	22	100.00	18.18	0.18	*Koyraboro Senni*†	4	26	0.00	0.00	—
Zoque (Copainala)	27	32	51.85	9.38	0.18	*Lezgian*†	8	27	0.00	3.70	—
Hawaiian†	10	57	10.00	1.75	0.18	*Murrinh-Patha*†	4	24	0.00	0.00	—
Hopi	23	40	86.96	15.00	0.17	Oksapmin†	16	35	0.00	0.00	—
German	36	30	97.22	16.67	0.17	*Rama*†	5	18	0.00	0.00	—
Mandarin	23	56	73.91	10.71	0.14	Sango†	30	60	0.00	0.00	—
Koiari	5	20	40.00	5.00	0.12	*Thai*†	10	54	0.00	1.85	—

Figure 5: Languages Sorted by Ratio of Markedness of Result Root (RR) Verbal Paradigms vs. PC Root Verbal Paradigms (the first two columns are the # of PC and result root paradigms where markedness could be determined, the next two the % of marked PC and result root verb paradigms, and the final column the ratio of % of marked result to marked PC root paradigms; boldface = more marked RR than PC root paradigms, italics = low data language, † = low verbal marking language, * = high marking language)

- Navajo also relies on equipollence, with most paradigms marked. However, there are a few suppletive verbs, slightly more among PC roots than result roots.
- Kinyarwanda has lots of labile relationships. Of the 24 PC paradigms only one was marked (4.17%) and of the 33 result paradigms only three were (9.09%). So the ratio is again a fluke.
- Kakataibo is also largely labile, but equipollence is a notable subpattern. The marked paradigms are entirely from this subclass, but are slightly imbalanced towards result roots, seen for 14 of 59 PC roots (23.73%) and 20 of 64 result roots (31.25%).

∴ In sum, the counterexamples are likely flukes of data plus language type.

#2 Seven languages have parity for PC and result root verbal markedness between 1.1 and .80.

- Five pattern like Navajo and Paumarí in relying on equipollence. In the 11 languages marked by a * 66.66% or more of all pairs within each paradigm are related equipollently (Fig. 8 in §C). Interestingly, all but one are found near the top of Fig. 5.

∴ The near perfect correlation with equipollent or underlying root systems and higher degrees of markedness regardless of root class is in fact expected: these languages tend to mark paradigm members across the board, thus obscuring the PC vs. result root distinction.

#3 Conversely, systematic non-derivational relationships explain other languages at parity. The 34 languages marked by a † have ≤33.33% marked paradigms among all roots.

- These include languages like Gooniyandi, Murrinh-Patha, Anejom̃, Sango, and Otomi that rely on labile or suppletive pairs for ≥66.66% relationships between all pairs within each paradigm, and languages that rely heavily on labile simple state/inchoative pairs (Burmese, Chukchi, Indonesian) or a mix of labile and suppletive relations (Barasano, Khoekhoe, Zulu).
- With little marking we might expect most to have a parity of 0.00 or incalculable, with some scattered throughout depending on how the small numbers work out, exactly as we see.

∴ Languages that generally do not rely on overt derivational relationships show low marking across both PC and result roots, again obscuring the distinction.

6 A Morphological Analysis

- Default realization rules for Asp and *v* heads might explain the morphological patterns. Functionally speaking, each category is marked for the opposite meaning, where $U_{Asp/v}$ refers to whatever is deemed an unmarked morphological form for the category in a given language and $M_{Asp/v}$ the marked form, assuming a markedness contrast exists:

(33) Default realization of v_{become} with complement $\sqrt{\text{ROOT}}$:
  a. If $\sqrt{\text{ROOT}}$ entails change, then U_v (result roots derive unmarked verbs)
  b. If $\sqrt{\text{ROOT}}$ does not entail change, then M_v (PC roots derive marked verbs)

(34) Default realization for Asp with complement X (root $\sqrt{\text{R}}$ or *v*P):
  a. If X does not entail change, then U_{Asp} (PC roots derive unmarked adjectives)
  b. If X entails change, then M_{Asp} (result roots derive marked adjectives)

- Equipollent languages may overtly mark everything and labile/suppletive languages nothing, i.e. in each case neutralizing the markedness distinction. In languages with markedness asymmetries the prediction is only the English type is attested, exactly as found.

- However, (33) go against some event structural approaches, where morphological rules should not be sensitive to root semantics. An alternative is to posit features that dictate the morphology. (e.g. as per Ramchand 2008, Alexiadou et al. 2006). However, this won't get the semantics, and once you analyze the semantics there's no need for the syntactic diacritics.

7 Conclusion and Extensions

- In sum, the PC/result root split recurs significantly across languages. Were it the result of some syntactic features of English there's no reason we'd see it so robustly preserved. But if it has to do with the semantic content of the roots themselves — that result roots entail change, contra the BTR — we should expect exactly this.

- The deeper question is *why* this split should exist, and why consistently across broad meaning category types. Beavers and Koontz-Garboden (2020) propose two conceptual explanations.

#1 There are states which are conceived of as occurring naturally ($\sqrt{}$FLAT), and another which as being the outcome of processes ($\sqrt{}$COOK). Naming such a state *requires* entailing a change.

(35) $[\![\sqrt{}\text{CRACK}]\!] = \lambda x \lambda s[cracked'(x,s)]$, where $\forall x \forall x[cracked'(x,s) \rightarrow \exists e'[become(e',s)]]$

#2 There are states that are so conventionally associated with coming about that it is functionally useful for to have roots lexicalizing the change with the state ($\sqrt{}$MELT; cp. $\sqrt{}$LIQUID).

(36) a. $[\![\sqrt{}\text{LIQUID}]\!] = \lambda x \lambda s[liquid'(x,s)]$ (*liquify again* can be restitutive)

b. $[\![\sqrt{}\text{MELT}]\!] = \lambda x \lambda s[liquid'(x,s) \wedge \exists e'[become(e',s)]]$ (*melt again* cannot)

- Either way, change will end up being part of root meanings, either by the nature of the concept or strong conventional associations. But what *else* can be in a root meaning?

- For some of these roots the actual process leading to change is open, predicting that under *again* modification it shouldn't matter what gave rise to the state in prior changes:

(37) [Mary hit the vase with a hammer, causing it to shatter into small pieces. She then painstakingly glued it back together. Afterwards she sang in an incredibly high-pitched voice at full volume, causing the vase to fall to pieces a second time.]
Mary shattered the vase again.

- But if there are states that must/conventionally arise from a process, are there also states that result from *particular* processes, or particular processes that must give rise to a state?

- Among change-of-state verbs a potential case would be verbs that supposedly violate so-called Manner/Result Complementarity (Rappaport Hovav and Levin 2010). For example, Beavers and Koontz-Garboden (2012) cite verbs of manner of cooking as such cases. We noted above that not just the state but the fact that the state changed must come about before.

- However, it seems that not just the change but the type of process must have occurred before:

(38) [Mary seared tofu and placed it in a steamer over boiling, flavored liquid, softening its texture. The tofu had an unfinished, uneven texture and was not cooked through. She seared it again and placed it in a covered pan in a hot oven in flavored liquid, evening out the texture and cooking it more thoroughly.]
#Mary braised the tofu again.

- This suggests the following denotation for something like $\sqrt{}$BRAISE, where crucially the templatic notion of causation is *also* a part of its meaning:

(39) $[\![\sqrt{}\text{BRAISE}]\!] = \lambda x \lambda s[braised'(x,s) \wedge \exists e' \exists v[cause'(v,e') \wedge become'(e',s) \wedge \forall v'[cause'(v',e') \rightarrow braising'(v')]]]$

- Focusing on the roots of canonical change-of-state verbs, we seem to have a typology:

(40) a. **Roots entailing a state**: $\sqrt{}$FLAT, $\sqrt{}$LONG, $\sqrt{}$WIDE, etc.
b. **Roots entailing a state that come about**: $\sqrt{}$CRACK, $\sqrt{}$BEND, etc.
c. **Roots entailing a state caused to come about**: $\sqrt{}$FRY, $\sqrt{}$BRAISE, etc.

- It looks like the roots typologize into classes that mimic major event templates: stative, inchoative, causative. So it's not quite that there's a free for all here. Rather, roots still fall into systematic types.

- These types also predict certain types of verbs in terms of grammatical behaviors:

(41)

	unmarked stative	**unmarked verb**	***again* reading**
$\sqrt{}$FLAT	Y	N	restitutive
$\sqrt{}$CRACK	N	Y	repetitive over change
$\sqrt{}$BRAISE	N	Y	repetitive over cause+change

- The expectation, though, is that we shouldn't have mixed properties — roots with "nonsensical" templatic meanings, or showing (say) simple states but high repetitive readings.

- Thus even if roots can mean a lot more things than bifurcation would have it, we at least still have a predictive theory of possible and impossible verbs.

- The broader picture then is that there are a range of types of roots that entail change, rooted in a cross-classification of how much additional templatic meaning they entail and what their syntax is. In any event, these data strongly argue against the BTR.

8 Acknowledgments

- This material is based upon work supported by the National Science Foundation under grant no. BCS-1451765. See https://verbal-roots.la.utexas.edu/. Henri Kauhanen was funded by Emil Aaltonen Foundation, The Ella and Georg Ehrnrooth Foundation, and the Federal Ministry of Education and Research (BMBF) and the Baden-Württemberg Ministry of Science as part of the Excellence Strategy of the German Federal and State Governments.

A Languages Used in the Typological Study

All macroareas and genetic affiliations are taken from WALS. Each language is listed with its genus and family, except where those are the same. Italicized languages are from WALS 200 and languages, languages not on WALS 200 or 100 are underlined, and data from boldfaced languages were collected or checked against a native speaker informants (for Spanish, German, and French data were collected from resources but checked by a native speaker). Otherwise all languages are from WALS 100 and data were collected through dictionary and grammatical resources .

Eurasia Basque (Basque)
Burmese (Burmese-Lolo, Sino-Tibetan)
Mandarin (Chinese, Sino-Tibetan)
Meithei (Kuki-Chin, Sino-Tibetan)
Burushaski (Burushaski)
Chukchi (Northern Chukotko-Kamchatkan, Chukotko-Kamchatkan)
English (Germanic, Indo-European)
German (Germanic, Indo-European)
Modern Greek (Greek, Indo-European)
Persian (Iranian, Indo-European)
Russian (Slavic, Indo-European)
Spanish (Romance, Indo-European)
French (Romance, Indo-European)
Hindi (Indic, Indo-European)
Finnish (Finnic, Uralic)
Georgian (Kartvelian)
Modern Hebrew (Semitic, Afro-Asiatic)
Japanese (Japanese)
Kannada (Southern Dravidian, Dravidian)
Khalkha (Mongolic, Altaic)
Korean (Korean)
Lezgian (Lezgic, Nakh-Daghestanian)
Thai (Kam-Tai, Tai-Kadai)
Turkish (Turkic, Altaic)
Vietnamese (Viet-Muong, Austro-Asiatic)

Africa Acholi (Nilotic, Eastern Sudanic)
Egyptian Arabic (Semitic, Afro-Asiatic)
Middle Atlas Berber (Berber, Afro-Asiatic)
Hausa (West Chadic, Afro-Asiatic)
Harar Oromo (Lowland East Cushitic, Afro-Asiatic)
Gújjolaay Eegimaa (Bak, Niger-Congo)
Swahili (Bantoid, Niger-Congo)
Kinyarwanda (Bantoid, Niger-Congo)
Zulu (Bantoid, Niger-Congo)
Sango (Ubangi, Niger-Congo)
Yoruba (Defoid, Niger-Congo)
Khoekhoe (Khoe-Kwadi)
Koyraboro Senni (Songhay)

Malagasy (Barito, Austronesian)

North America
Plains Cree (Algonquian, Algic)
Hopi (Hopi, Uto-Aztecan)
Yaqui (Cahita, Uto-Aztecan)
Jakaltek (Mayan)
Tenango Tzeltal (Mayan)
Karok (Karok)
Kiowa (Kiowa-Tanoan)
Koasati (Muskogean)
Lakhota (Core Siouan, Siouan)
Chalcatongo Mixtec (Mixtecan, Oto-Manguean)
Mezquital Otomí (Otomian, Oto-Manguean)
Navajo (Athapaskan, Na-Dene)
Oneida (Northern Iroquoian, Iroquoian)
Rama (Rama, Chibchan)
Tsimshian (Penutian)
Yup'ik (Eskimo, Eskimo-Aleut)
Zoque (Mixe-Zoque)

South America
Barasano (Tucanoan)
Carib (Cariban)
Guaraní (Tupi-Guaraní, Tupian)
Minica Huitoto (Huitoto, Huitotoan)
Kakataibo (Cashibo-Cacataibo, Panoan)
Mapudungun/Mapuche (Araucanian)
Mocoví (South Guaicuran, Guaicuran)
Paumarí (Arauan)
Huallaga Quechua (Quechuan)
Warao (Warao)
Yagua (Peba-Yaguan)

Papunesia
Alamblak (Sepik Hill, Sepik)
Kwoma (Middle Sepik, Sepik)
Anejom̃ (Oceanic, Austronesian)
Bariai (Oceanic, Austronesian)
Fijian (Oceanic, Austronesian)
Hawaiian (Oceanic, Austronesian)
Chamorro (Chamorro, Austronesian)
Indonesian (Malayo-Sumbawan, Austronesian)
Paiwan (Paiwan, Austronesian)
Tagalog (Greater Central Philippine, Austronesian)
Lower Grand Valley Dani (Dani, Trans-New Guinea)
Kewa (Engan, Trans-New Guinea)
Koiari (Koiarian, Trans-New Guinea)
Daga (Dagan)
Oksapmin (Oksapmin)

Australia Gooniyandi (Bunuban)
Kayardild (Tangkic, Tangkic)
Martuthunira (Western Pama-Nyungan, Pama-Nyungan)
Pintupi (Western Pama-Nyungan, Pama-Nyungan)
Murrinh-Patha (Murrinh-Patha, Southern Daly)
Tiwi (Tiwian)

B Full Paradigm for Kinyarwanda Roots

Note that the simple stative form is also used to express the result state. We took this to mean it had a simple state meaning but in context it was clear there this state was a result state. There are no underlying roots in this language; we omit this column for space reasons.

Root	Simple State	Inchoative	Causative	Result State
large/big/enlarge	/nini	kw-iyongera	k-ongera	
small/shrunk/shrink	/toya		ku-gabanya	
short/shorten	/gufi	kw-ihina	ku-gabanya	
long/lengthen	/re-re	kw-iyongera	k-ongera	
deep/deepen	/re-re	gu-tebera	k-ongera	
wide/widen	/gari	kw-iyongera	k-ongera	
tall/height/heighten	/-re-re	ku-bara -re-re	k-ongera	
aged/age	gu-saza	gu-saza		
bad/worse/worsen	/bi	ku-bara mbikurushaho		
good/improved/improve	/iza	kw-iyongera		
white/whiten	kw-era			
white/whiten	umweru	gu-hinduka umweru		
black/blacken	umukara	gu-hinduka umukara		
red/redden	umutuku	gu-hinduka umutuku		
green/make green	icyatsi	gu-hinduka icyatsi		
blue/make blue	ubururu	gu-hinduka ibururu		
brown/make brown	ikaki	gu-hinduka igitaka		
cold/make cold	gu-konja	gu-konja	gu-konjesha	
hot/heat up	gu-shyuha	gu-shyuha	gu-shyushya	
dirty/dirty	umwanda			
dirty/dirty	kw-andura	kw-andura	kw-anduza	
dry/dry	ku-ma	ku-ma	ku-misha	
wet/wetten	gu-toha	gu-toha	gu-tosa	
straight/straighten	ku-gororoka		ku-gorora	
hard/harden	gu-komera	gu-komera		
soft/soften	k-oroha	k-oroha	k-oroheza	
clean/clean	gu-sukura	gu-sukura	gu-sukura	
sharp/sharpen (a point)	gu-songora		gu-songora	
sharp/sharpen	gu-tyara	gu-tyara/gu-tyaza	gu-tyaza	
sweet/sweeten	ku-ryoha	ku-ryoha	ku-ryoheza	
strong/strengthen	gu-komera	gu-komera	gu-komeza	
fast/speed up	kw-ihuta	k-ongera umuvuduko	kw-ihutisha	
slow/slow down	buhoro	ku-gabanya umuvuduko		
angry/anger	ku-rakara	ku-rakara	ku-rakaza	
calm/calmed	kw-ikaruma	kw-ikaruma	gu-karumisha	
frightened/frighten	gu-tinya	gu-tinya	gu-tinyisha	

sick/sicken	ku-rwara	ku-rwara	
sad/sadden	ku-babara	ku-babara	ku-babaza
hurt/hurt	ku-babara	ku-babara	ku-babaza
tired/tire	ku-nanirwa	ku-nanirwa	ku-naniza
entertained/entertain	ku-nezerwa	ku-nezerwa	ku-nezeza
surprised/surprise	gu-tungurwa	gu-tungurwa	gu-tungura
worried/worry	ubwoba		
pleased/please	kw-ishima	kw-ishima	gu-shimisha

English Gloss	Simple State	Inchoative	Causative	Result State
burned/burn		gu-shiririra	gu-shiririza	gu-shiririra
melted/melt		gu-shonga	gu-shongesha	
congeal		gu-fata	gu-fatisha	gu-fata
decayed/decay, rotten/rot		ku-bora	ku-boza	
swollen/swell			ku-byimbisha	ku-byimba
grown/grow		gu-kura	gu-kuza	
bloomed/bloom		ku-bumbura		ku-bumbura
wilted/wilt		ku-raba	ku-rabiza	ku-raba
withered/wither		ku-raba	ku-rabiza	ku-namba
sprouted/sprout		gu-tunguka		gu-tunguka
tarnished/tarnish		gu-koboka		
roast/roasted			k-otsa	
cook/cooked			gu-teka	gu-tek-ek-a
fry/fried		gu-karang-w-a	gu-karanga	gu-karang-ik-a
broken/break		ku-men-ek-a	ku-mena	ku-men-ek-a
cracked/crack		gu-saduka	gu-satura	
crushed/crush		gu-hond-ek-a	gu-honda	gu-hond-ek-a
shattered/shatter		gu-shwanyuka	gu-shwanyura	
torn/tear, ripped/rip		gu-c-ik-a	gu-ca	gu-c-ik-a
snapped/snap		ku-vun-ik-a	ku-vuna	ku-vun-ik-a
bent/bend		ku-gonda	ku-gonda	ku-gond-ek-a
folded/fold			gu-hina	
wrinkled/wrinkle		kw-ihinarika	gu-hinarika	
dead/killed/kill		gu-pfa	kw-ica	gu-pfa
drowned/drown		ku-rohama	kw-ibiza	ku-rohama
destroyed/destroy, ruined/ruin		kw-angir-ik-a	kw-angiza	kw-angir-ik-a
go up		ku-zamuka	ku-zamura	
go down		ku-manuka	ku-manura	
differing/differ		gu-tandukanya	gu-tandukanya	
come/came		ku-za	ku-zana	
gone/go		ku-genda	ku-jyana	
go in (entered/enter)		kw-injira	kw-injiza	
go out (exited/exit)		gu-sohoka	gu-sohora	
returned/return		gu-subira	gu-subiza	

C Numerical Data for Typological Study

PC root	#states	#languages	attested	Result root	#states	#languages	attested
aged/age	0	81	0.00%	bent/bend	6	73	8.22%
bad/worse/worsen	80	84	95.24%	bloomed/bloom, flowered/flower, blossomed/blossom	4	65	6.15%
black/blacken	83	84	98.81%	boiled/boil	2	77	2.60%
blue/make blue	66	67	98.51%	broken/break	1	85	1.18%
brown/make brown	54	55	98.18%	burned/burn	3	82	3.66%
clean/clean	54	67	80.60%	come/came	0	81	0.00%
clear/clear	50	57	87.72%	cooked/cook, baked/bake, fried/fry, roasted/roast, steamed/steam	0	86	0.00%
cold/make cold	83	83	100.00%	cracked/crack	1	63	1.59%
cool/cool	54	63	85.71%	crushed/crush	0	71	0.00%
deep/deepen	71	72	98.61%	dead/killed/kill	5	87	5.75%
dirty/dirty	74	78	94.87%	decayed/decay, rotten/rot	2	79	2.53%
dry/dry	72	85	84.71%	destroyed/destroy, ruined/ruin	0	70	0.00%
fast/speed up	63	71	88.73%	differing/differ	40	52	76.92%
good/improved/improve	83	85	97.65%	drowned/drown	1	71	1.41%
green/make green	71	72	98.61%	fermented/ferment	3	50	6.00%
hard/harden, tough/toughen	74	79	93.67%	folded/fold	0	64	0.00%
hot/heat up	80	83	96.39%	frozen/freeze	5	42	11.90%
large/big/enlarge	86	87	98.85%	go down (fallen/fall, dropped/drop, descended/descend, decreased/decrease, declined/decline)	1	85	1.18%
long/lengthen	80	82	97.56%	go in (entered/enter)	0	76	0.00%
red/redden	77	80	96.25%	go out (exited/exit)	0	63	0.00%
sharp/sharpen	67	75	89.33%	go up (raised/rise, ascended/ascend, increased/increase, gained/gain)	2	83	2.41%
short/shorten	76	77	98.70%	gone/go	0	78	0.00%
slow/slow down	63	66	95.45%	grown/grow	3	70	4.29%
small/shrunk/shrink	81	84	96.43%	melted/melt	3	64	4.69%
smooth/smooth	69	73	94.52%	murdered/murder	0	45	0.00%
soft/soften	69	71	97.18%	returned/return	0	72	0.00%
straight/straighten	71	76	93.42%	rusted/rust	10	53	18.87%
strong/strengthen	76	80	95.00%	shattered/shatter	1	53	1.89%
sweet/sweeten	71	72	98.61%	snapped/snap	0	39	0.00%
tall/height/heighten	70	73	95.89%	split/split	0	67	0.00%
tight/tighten	55	69	79.71%	sprouted/sprout, germinated/germinate	1	63	1.59%
warm/warm	63	68	92.65%	swollen/swell	1	79	1.27%
weak/weaken	61	68	89.71%	tarnished/tarnish	3	32	9.38%
wet/wetten	71	83	85.54%	torn/tear, ripped/rip	0	77	0.00%
white/whiten	81	84	96.43%	withered/wither, wilted/wilt	1	59	1.69%
wide/widen	76	78	97.44%	wrinkled/wrinkle, creased/crease	2	61	3.28%

Figure 6: Percentage of Languages with a Simple State for a Given Root when there is Data for that Root (#States = Number of languages with a simple state, #Languages = Number of languages with any data for that root, Attested = #Roots/#Languages)

PC root	#marked	#languages	marked	Result root	#marked	#languages	marked
aged/age	10	36	27.78%	bent/bend	14	57	24.56%
bad/worse/worsen	21	39	53.85%	bloomed/bloom, flowered/flower, blossomed/blossom	5	54	9.26%
black/blacken	27	35	77.14%	boiled/boil	13	72	18.06%
blue/make blue	10	15	66.67%	broken/break	21	80	26.25%
brown/make brown	10	12	83.33%	burned/burn	11	79	13.92%
clean/clean	14	38	36.84%	come/came	4	80	5.00%
clear/clear	13	37	35.14%	cooked/cook, baked/bake, fried/fry, roasted/roast, steamed/steam	12	79	15.19%
cold/make cold	27	45	60.00%	cracked/crack	11	55	20.00%
cool/cool	30	44	68.18%	crushed/crush	6	64	9.38%
deep/deepen	17	21	80.95%	dead/killed/kill	9	86	10.47%
dirty/dirty	21	42	50.00%	decayed/decay, rotten/rot	11	58	18.97%
dry/dry	32	64	50.00%	destroyed/destroy, ruined/ruin	9	64	14.06%
fast/speed up	15	33	45.45%	differing/differ	8	27	29.63%
good/improved/improve	30	46	65.22%	drowned/drown	12	68	17.65%
green/make green	14	16	87.50%	fermented/ferment	4	43	9.30%
hard/harden, tough/toughen	26	43	60.47%	folded/fold	6	53	11.32%
hot/heat up	28	54	51.85%	frozen/freeze	5	38	13.16%
large/big/enlarge	25	47	53.19%	go down (fallen/fall, dropped/drop, descended/descend, decreased/decrease, declined/decline)	16	83	19.28%
long/lengthen	26	37	70.27%	go in (entered/enter)	6	71	8.45%
red/redden	31	35	88.57%	go out (exited/exit)	7	60	11.67%
sharp/sharpen	14	46	30.43%	go up (raised/rise, ascended/ascend, increased/increase, gained/gain)	16	82	19.51%
short/shorten	21	35	60.00%	gone/go	4	73	5.48%
slow/slow down	12	23	52.17%	grown/grow	9	67	13.43%
small/shrunk/shrink	17	51	33.33%	melted/melt	16	61	26.23%
smooth/smooth	14	32	43.75%	murdered/murder	5	41	12.20%
soft/soften	21	37	56.76%	returned/return	11	70	15.71%
straight/straighten	17	38	44.74%	rusted/rust	9	37	24.32%
strong/strengthen	21	38	55.26%	shattered/shatter	7	48	14.58%
sweet/sweeten	12	21	57.14%	snapped/snap	4	37	10.81%
tall/height/heighten	15	31	48.39%	split/split	11	60	18.33%
tight/tighten	17	39	43.59%	sprouted/sprout, germinated/germinate	5	59	8.47%
warm/warm	26	45	57.78%	swollen/swell	13	65	20.00%
weak/weaken	24	37	64.86%	tarnished/tarnish	9	30	30.00%
wet/wetten	19	52	36.54%	torn/tear, ripped/rip	16	70	22.86%
white/whiten	27	35	77.14%	withered/wither, wilted/wilt	7	46	15.22%
wide/widen	22	35	62.86%	wrinkled/wrinkle, creased/crease	15	43	34.88%

Figure 7: Percentage of Languages with a Marked Verbal Paradigm for a Given Root when there is Sufficient Data to Determine Markedness (#Marked = Number of languages with a marked paradigm, #Languages = Number of languages sufficient data, Marked = #Marked/#Languages)

Language	# Roots	% Und	# Pairs	%Equi	% ND	Language	# Roots	% Und	# Pairs	%Equi	% ND
Acholi	102	51.96	128	48.05	50.78	Koasati	70	0.00	67	0.75	17.91
Alamblak	20	10.00	10	0.00	10.00	Koiari	57	15.79	48	22.92	31.25
Anejom̃	85	0.00	54	0.00	88.89	Korean	94	11.70	88	27.27	34.09
Arabic (Egyptian)	76	93.42	272	45.77	27.76	Koyraboro Senni	55	0.00	54	11.11	9.26
Barasano	73	0.00	45	4.44	66.67	Kwoma	63	0.00	36	0.00	69.44
Bariai	85	1.18	37	2.70	37.84	Lakhota	68	2.94	70	12.14	23.57
Basque	82	9.76	147	10.88	45.58	Lezgian	81	3.70	29	6.90	44.83
Berber (Middle Atlas)	94	67.02	187	8.56	32.09	Malagasy	73	53.42	119	66.39	5.88
Burmese	88	1.14	90	1.11	67.22	Mandarin	101	3.96	190	13.16	38.68
Burushaski	106	10.38	92	4.35	38.04	Mapudungun	75	2.67	85	4.71	30.59
Carib	48	52.08	47	55.32	14.89	Martuthunira	50	0.00	36	19.44	11.11
Chamorro	77	0.00	39	2.56	46.15	Meithei	86	4.65	61	14.75	19.67
Chukchi	51	0.00	22	4.55	25.00	Mixtec (Chalcatongo)	93	5.38	53	16.98	11.32
Cree (Plains)	105	19.05	171	71.35	22.51	Mocovi	70	4.29	120	7.92	36.25
Daga	77	0.00	23	4.35	43.48	Murrinh-Patha	49	0.00	29	0.00	96.55
Dani (Lower Grand Valley)	31	29.03	29	44.83	6.90	Navajo	59	100.00	100	44.00	34.00
English	103	0.00	432	3.01	37.50	Oksapmin	62	0.00	15	0.00	80.00
Fijian	66	3.03	63	15.87	4.76	Oneida	55	98.18	57	58.77	34.21
Finnish	104	15.38	469	21.64	23.99	Oromo (Harar)	57	28.07	88	94.89	4.55
French	87	0.00	377	11.27	18.17	Otomi	172	2.33	97	11.34	67.01
Georgian	78	5.13	38	52.63	10.53	Paiwan	60	61.67	71	52.11	9.86
German	68	11.76	287	14.29	17.60	Paumar	61	90.16	52	75.00	15.38
Gooniyandi	33	0.00	3	0.00	100.00	Persian	91	14.29	100	15.00	30.00
Greek (Modern)	76	2.63	154	8.44	42.86	Pintupi	75	21.33	73	31.51	6.85
Guaran	95	1.05	95	4.21	46.32	Quechua (Huallaga)	93	9.68	62	9.68	14.52
Gjjolaay Eegimaa	75	97.33	263	5.70	47.91	Rama	54	0.00	40	15.00	5.00
Hausa	73	20.55	90	50.56	26.67	Russian	82	4.88	273	15.02	20.51
Hawaiian	116	0.86	88	3.41	39.77	Sango	102	3.92	171	4.09	87.13
Hebrew (Modern)	89	100.00	315	53.02	25.87	Spanish	74	2.70	287	17.60	8.01
Hindi	93	2.15	172	22.97	19.77	Swahili	98	20.41	177	22.60	22.32
Hopi	75	21.33	142	35.21	5.63	Tagalog	93	38.71	78	69.23	19.23
Huitoto (Minica)	85	94.12	46	27.17	63.04	Tenango Tzeltal	76	5.26	303	19.80	10.23
Indonesian	94	15.96	159	3.77	37.42	Thai	102	3.92	94	3.19	60.64
Jakaltek	57	0.00	16	0.00	43.75	Tiwi	45	0.00	17	5.88	52.94
Japanese	83	2.41	100	38.00	20.00	Tsimshian (Coast)	73	6.85	38	2.63	65.79
Kakataibo	123	26.02	531	14.69	16.57	Turkish	89	8.99	199	14.32	5.28
Kannada	71	14.08	30	13.33	50.00	Vietnamese	71	0.00	144	18.06	22.92
Karok	75	13.33	29	24.14	27.59	Warao	60	1.67	104	15.87	22.60
Kayardild	49	0.00	45	26.67	0.00	Yagua	124	4.84	151	9.93	47.68
Kewa	71	0.00	13	0.00	76.92	Yaqui	76	1.32	85	26.47	22.94
Khalkha	65	13.85	167	27.25	8.68	Yoruba	175	2.29	204	2.94	62.75
Khoekhoe	89	0.00	124	7.26	28.23	Yup'ik	66	12.12	88	13.64	20.45
Kinyarwanda	69	5.80	113	3.54	44.25	Zoque (Copainala)	69	0.00	149	14.09	14.43
Kiowa	23	0.00	9	11.11	0.00	Zulu	84	0.00	212	16.75	39.15

Figure 8: Data on Language Types (#Roots = Number of roots with any data, %Und = Percentage of roots with data that have an underlying root, #Pairs = Number of related pairs in data set, %Equi = Percentage of pairs related equipollently, %ND = Percentage of pairs related non-derivationally, i.e. as labile or suppletive)

References

Alexiadou, Artemis, Elena Anagnostopoulou, and Florian Schäfer. 2006. The properties of anti-causatives crosslinguistically. In M. Frascarelli, ed., *Phases of Interpretation*, pages 187–211. Berlin: Mouton de Gruyter.

Arad, Maya. 2005. *Roots and Patterns: Hebrew Morpho-Syntax*. Dordrecht: Springer.

Beavers, John and Andrew Koontz-Garboden. 2012. Manner and result in the roots of verbal meaning. *Linguistic Inquiry* 43:331–369.

Beavers, John and Andrew Koontz-Garboden. 2020. The Roots of Verbal Meaning. Oxford: Oxford University Press.

Borer, Hagit. 2005. *Structuring Sense, Vol. I and II*. Oxford: Oxford University Press.

Deo, Ashwini, Itamar Francez, and Andrew Koontz-Garboden. 2011. The morphosemantics of *–ed*. Paper presented at Dimensions of grammar in honor of Paul Kiparsky. Konstanz, August 3, 2011.

Deo, Ashwini, Itamar Francez, and Andrew Koontz-Garboden. 2013. From change to value difference in degree achievements. In *Proceedings of Semantics and Linguistic Theory 23*, pages 97–115. Ithaca, NY: CLC Publications.

Dixon, R.M.W. 1982. *Where have all the adjectives gone?: And other essays in semantics and syntax*. The Hague: Mouton.

Dowty, David. 1979. *Word meaning and Montague grammar*. Dordrecht: D. Reidel Publishing.
Dryer, Matthew S. and Martin Haspelmath, eds. 2013. *WALS Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
Dunbar, Ewan and Alexis Wellwood. 2016. Addressing the "two interface" problem: Comparatives and superlatives. *Glossa* 1(1):1–29.
Embick, David. 2004. On the structure of resultative participles in English. *Linguistic Inquiry* 35(3):355–392.
Embick, David. 2009. Roots, states, and stative passives. Talk given that the 2009 Roots Workshop, Stuttgart.
Gawron, Mark. 2009. The lexical semantics of extent verbs. San Diego State University, January 31, 2009.
Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago, IL: University of Chicago Press.
Harley, Heidi. 2008. On the causative construction. In S. Miyagawa and M. Saito, eds., *The Handbook of Japanese Linguistics*, pages 20–53. Oxford: Oxford University Press.
Jerro, Kyle. 2017. Change-of-state verb roots in Kinyarwanda. Paper presented at the Linguistic Society of America 91st Annual Meeting, Austin, TX.
Koontz-Garboden, Andrew. 2005. On the typology of state/change of state alternations. *Yearbook of Morphology* 2005:83–117.
Koontz-Garboden, Andrew. 2006. The states in changes of state. Proceedings of the 32nd annual meeting of the Berkeley Linguistics Society.
Koontz-Garboden, Andrew. 2010. The lexical semantics of derived statives. *Linguistics and Philosophy* 33:285–324.
Lakoff, George. 1965. *On the nature of syntactic irregularity*. Ph.D. thesis, Indiana University, Bloomington, IN. Published 1970 as *Irregularity in syntax*. New York: Holt, Rinehart, and Winston.
Levin, Beth. 1993. *English verb classes and alternations: A preliminary investigation*. Chicago, IL: University of Chicago Press.
Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In *Proceedings of the 21st Annual Penn Linguistics Colloquium*, vol. 4.2 of *University of Pennsylvania Working Papers in Linguistics*, pages 201–225.
Megerdoomian, Karine. 2002. *Beyond words and phrases: A unified theory of predicate composition*. Ph.D. thesis, University of Southern California, Los Angeles, CA.
Nedjalkov, Vladimir P. and Sergej Je. Jaxontov. 1988. The typology of resultative constructions. In V. P. Nedjalkov, ed., *Typology of resultative constructions*, pages 3–62. Amsterdam: John Benjamins.
Pross, Tillmann. 2016. What about word meaning if syntax is the only generative component of languages? Talk given at UT Austin, May 19.
Ramchand, Gillian. 2008. *Verb Meaning and the Lexicon: A First Phase Syntax*. Cambridge: Cambridge University Press.
Rappaport Hovav, Malka. 2010. Building scalar changes. The Hebrew University of Jerusalem.
Rappaport Hovav, Malka and Beth Levin. 1998. Building verb meanings. In M. Butt and W. Geuder, eds., *The Projection of Arguments: Lexical and Compositional Factors*, pages 97–133. Stanford: CSLI Publications.
Rappaport Hovav, Malka and Beth Levin. 2010. Reflections on manner/result complementarity. In E. Doron, M. Rappaport Hovav, and I. Sichel, eds., *Syntax, Lexical Semantics, and Event*

Structure, pages 21–38. Oxford: Oxford University Press.
Shibatani, Masayoshi. 1976. The grammar of causative constructions: A conspectus. In M. Shibatani, ed., *The Grammar of Causative Constructions*, pages 1–39. New York: Academic Press.
Shibatani, Masayoshi and Prashant Pardeshi. 2001. The causative continuum. In M. Shibatani, ed., *The Grammar of Causation and Interpersonal Manipulation*, pages 85–126. Amsterdam: John Benjamins.
Spathas, Giorgos. 2017. Results in the decomposition of verbal predicates: Evidence from additive operators. Talk given at The Building Blocks and Mortar of Meaning II, University of Tübingen, 3-4 November.
Valle, Daniel, John Beavers, and Andrew Koontz-Garbodden. 2017. States and change-of-state in kakataibo: The role of the root. Paper to be presented at the Society for the Study of the Indigenous Languages of the Americas.
Watkins, Laurel J. and Parker McKenzie. 1984. *A grammar of Kiowa*. Studies in the Anthropology of North American Indians. Lincoln / London: University of Nebraska Press.
Wechsler, Stephen. 2005. What is right and wrong about little-*v*. In M. Vulchanova and T. A. Åfarli, eds., *Grammar and Beyond — Essays in Honour of Lars Hellan*, pages 179–195. Oslo: Novus Press.